Predictions, precision, and agentive attention

Citation for published version:
Clark, A 2018, 'Predictions, precision, and agentive attention', *Consciousness and Cognition*, vol. 56, pp. 115-119. https://doi.org/10.1016/j.concog.2017.06.013

Digital Object Identifier (DOI):
10.1016/j.concog.2017.06.013

Published In:
Consciousness and Cognition

Predictions, Precision, and Agentive Attention

Forthcoming in *Consciousness and Cognition*

Corresponding Author:

Andy Clark
School of Philosophy, Psychology, and Language Sciences
University of Edinburgh,
Edinburgh, UK
EH12 5AY

Abstract

Ransom, Fazelpour, and Mole (this journal - 2017) raise an important puzzle for the 'prediction error minimization' account of cognitive processing. That account depicts all cognitive processing as fundamentally in the business of minimizing prediction errors concerning the evolving flow of sensory information. One of the cornerstones of these highly ambitious, would-be unifying accounts is their depiction of attention as *nothing other than* the process of optimizing the precision (inverse variance) of critical prediction error signals. But that story, Ransom et al suggest, cannot accommodate voluntary shifts of attention. In this paper, I show why this challenge to the grand unifying project fails. It fails because it locates the origins of voluntary attention in complexes of unanalyzed desire rather than in changing complexes of beliefs.

Keywords: Attention; voluntary attention; prediction; prediction error minimization; mental action.

Predictions, Precision, and Agentive Attention

Abstract

Ransom, Fazelpour, and Mole (this journal - 2017) raise an important puzzle for the 'prediction error minimization' account of cognitive processing. That account depicts all cognitive processing as fundamentally in the business of minimizing prediction errors concerning the evolving flow of sensory information. One of the cornerstones of these highly ambitious, would-be unifying accounts is their depiction of attention as *nothing other than* the process of optimizing the precision (inverse variance) of critical prediction error signals. But that story, Ransom et al suggest, cannot accommodate voluntary shifts of attention. In this paper, I show why this challenge to the grand unifying project fails. It fails because it locates the origins of voluntary attention in complexes of unanalyzed desire rather than in changing complexes of beliefs.

1. A Model of Attention

Cognition, it has recently been claimed, is always and everywhere a matter of minimizing prediction errors concerning the evolving flow of sensory information (Friston (2005, 2010), Hohwy (2013), Clark (2013a, 2016)). Such minimization occurs within a complex, multi-area, multi-level processing regime whose hierarchical structure reflects the operation of interacting and deeply nested environmental causes. But for all this to work, such systems need to be sensitive not simply to the presence of specific prediction errors, but to their state-dependent (context-variable) significance. Only by being thus sensitive can intelligent systems vary the impact of specific error signals according to their task-salience, reliability, or estimated value. This is the key role played by the so-called 'precision-weighting' of prediction error. Precision-weighting reflects systemic estimates of the inverse variance of specific prediction error signals. More precise (lower variance) signals receive higher weighting, and are correspondingly positioned to exert greater influence over subsequent processing. In such cases variable precision weighting turns up the (post-synaptic) 'volume' on the selected signals.

Attention, as Ransom et al (2017) rightly stress, is then identified with this mechanism (which may have multiple implementations in the brain, operating in different ways and at different timescales – see Clark (2013b)). Most generally, it is claimed that "attention is simply the process of optimizing precision during hierarchical inference" (Friston (2009) p.299). Prediction errors are thus re-scaled, courtesy of recurrent connections and other devices, so as to increase or decrease their influence as task and context demand.

The ability of precision variation to enforce decreased influence is especially important in cases of self-generated movement. Here the precision of information specifying the *current* state of the motor-effector system is attenuated, allowing top-down predicted states and trajectories to prevail. Precise prediction errors representing the difference between predicted and actual states are then minimized, bringing the motor plant into line with top-down predictions. Effective motor control thus requires a kind of systematic *dis-attention* enabling motor activity to be guided by the agent's intentions (Friston et al (2011), Adams et al (2013), Brown et al (2011), Brown, Adams et al (2013)).

In sum, these stories depict attention and sensory attenuation as two sides of the same coin, mediated by augmenting and attenuating sensory precision respectively. Variable precision-weighting is thus revealed as a powerful control mechanism whose role goes far beyond the simple estimation of signal-to-noise.

2. Attention Switching

Ransom et al (2017) argue that this picture cannot do justice to the full range of ways in which attention is allocated. To demonstrate this, they focus upon the selective direction of attention, using Neisser and Becklen's famous (1975) study as a key resource. In this study, subjects are exposed to two different films presented simultaneously in a single part of their visual field. This is an ecologically abnormal situation, supported experimentally by the use of half-silvered mirrors which create an effect "that is something like the effect of looking out through the window of a lit room at dusk, when the exterior and reflected interior worlds are both visible in the same part of space." (Ransom et al (2017) p.102). The superimposed films depicted different events, unfolding at different distances. But each film depicted the playing of a game – in one case a hand-clapping game, filmed up-close, and in the other case a ball game, filmed from a greater distance. Sometimes both film-streams were shown to each eye, and at other times each eye received a single (but different) film-stream. The results were that, when both scenes were shown to each eye, subjects had no trouble selectively attending to one stream, and could easily switch streams on demand. Attentive perception of both streams at once was, however, difficult or impossible.

At first pass, nothing here is problematic for PEM. At any given moment, a specific hypothesis ('hand-clapping' or 'ball-game') is used to meet the incoming sensory signal, and high-precision assigned to errors conditioned on that hypothesis. So we bring that sensory information into clear focus, while down-grading the precision on those aspects of the signal that would otherwise be recruiting the alternative hypothesis. There is no combined hypothesis that has any plausibility given prior learning, so we cannot attend to (assign high-precision to errors concerning) both at once.

The trouble – according to Ransom et al – comes when we try to explain why subjects find it so easy to switch between the streams. For the switch need be consequent upon no external cue. Subjects are able to flip between the streams 'at will', without any change in the incoming signal or elsewhere in the sensorium. These easy flips of focus are not consequent upon any changes whatsoever in the incoming signals but are rather "directed by mental occurrences that are internal to the attentive perceiver" (op cit. p. 104). Those occurrences might be decisions, or emotion-led (perhaps we are bored with one stream), and so on.

The core puzzle can now be brought into view. First, there is nothing in the external signal that marks out one stream of information as more reliable, valuable, or precise than the other. So there seems to be no external cue that could prompt the system to set the gain higher on one set of signals. Second, many of the thoughts that might prompt the flip have *non-indicative contents*. They are not about the way the world is, but might instead consist in something like the desire to look at the second stream. But it is not obvious (though more on this later) that non-indicative contents can alter precision-weightings, which are meant to track the reliability of signals rather than e.g. their desirability. Pushing this line of argument, Ransom et al note that "the precision of an incoming signal is a statistical feature of that signal, over some interval" (op cit p.106). The trouble, then, is that "the precisions *in the world* are equal" (op cit p.106, italics in original).

3. The Real Puzzle – Mental Action

Thus expressed, the puzzle may seem to have a simple solution. The solution is that although the incoming signals are in fact equally reliable and clear, the agent simply starts to treat one set as stronger or more reliable. Cast in PP terms, that means she starts to expect that signal to be stronger or more reliable, increases the gain on it, and thus brings about that very effect. Predictions of reliability/strength impact precision-weightings, bringing it about that one set are indeed stronger, clearer, better signals. This is because, by upping the gain on just the right aspects of the signal, that signal does in fact become better differentiated from background noise, hence (in the technical, 'inverse variance' sense) more precise. The situation, in the case at hand, would be roughly akin to having two radios, tuned to different channels, and turning up the volume on one of them. Increasing the estimated precision is like turning up the volume, and hence results in making that signal louder. The selective increase in precision thus becomes a kind of 'self-fulfilling prophecy' – estimating greater precision brings greater precision about.

Ransom et al consider this move, but reject it by arguing that (in the specific case at hand) it cannot explain the agent's ability to shift her focus at will from one stream to the other. This, they argue, is because the self-fulfilling prophecy should now become locked in place, preventing the easy flips that are in fact observed. Insofar as the first estimate really does make it the case that the signals from one stream are more precise, that stream should continue to dominate.

What is missing, it now seems clear, is a compelling account of the *agentive (*or as they put it, 'voluntary') *re-allocation* of precision. The worry that Ransom et al are pressing depends on the picture of precision as simply estimating signal reliability. If we artificially inflate the reliability estimate for stream one (say), we create a situation in which those signals are rendered more precise, enjoying greater post-synaptic impact. But this move, they argue, leaves us with no apparatus *internal to the PP story* capable of explaining how those precision estimations may endogenously alter. Even if we grant that precision estimations can act as a kind of 'self fulfilling prophecy', there is (so the worry goes) no story about why precision estimations alter in just those ways at just those moments (when voluntary attention shifts). In fact, it starts to look as if some kind of voluntary attentional shift must precede (and drive) the selection of the new materials to which the 'self fulfilling prophecy trick' then applies.

The worry, in other words, is that we are left with no PP story about the top-down distribution of attention itself. PP, if their argument is right, at best explains how endogenous attending is accomplished (by increasing precision) but not how it is determined moment-by-moment, becoming allocated and re-allocated as a result of decisions, motivations, and other mental actions. But an account of attention that does not encompass the top-down selectivity of attention is radically incomplete. More generally, the worry is that a PP model of cognition cannot accommodate mental actions and hence is severely incomplete. It is beyond the scope of this reply to address the full gamut of mental actions. But I think we can make progress, and begin to gesture at the shape of such a larger story, by looking harder at the kind of case that Ransom et al have here brought into clear focus. It is to this task that we next turn.

4. Desires, Predictions, and Beliefs

PEM claims that agentive attention consists in the allocation and re-allocation of precision, in ways that vary according to task and context. When external circumstances vary, precision estimations may vary accordingly. As the fog increases, the estimated precision of the visual signal goes down. Changes of task have the same kind of effect. If, looking at my desk, I want to find the key ring, the estimated precision of signal elements that would indicate a shiny, silver-colored object is increased. If I want to find a certain paper, the estimated precision for signal elements relevant to that paper (it's look, size, or even title) increase. Indeed, there are fairly detailed accounts of the neurobiology that would mediate just this kind of top-down control over precision, for example in figure-ground segregation (Kanai et al 2015).

On the face of it, the Neissser and Becklen study has exactly that form. To see the ball-game, increase precision on ball-game-y aspects of the overall signal. To see the hand-clapping game, increase the precision on hand-clappy stuff. The puzzle that Ransom et al pose does not bite here. Rather, it aims to cast doubt on PP's capacity to explain how we recruit those shifts in precision estimation. They cannot be recruited by those changing estimates of precision, since they *are* those changing estimates of precision. They cannot be recruited by changing external cues, since the external cues have not changed. If the PEM account of attention is all about precision estimation, and precision estimation is all about reliability estimation, it is hard to see what further resources may be brought to bear from within the PEM story about attention itself. So PEM offers an incomplete account of endogenous selective attention. Intuitively, one wants to say that the altered precision estimations are consequent upon changing internal cues, and that these are indeed part and parcel of a PEM account of selective attention. I think this is broadly speaking correct, but it needs careful unpacking to avoid the kinds of worry that Ransom et al are raising.

It is important to notice that the PEM model of attention *allocation* (see also Hohwy (2013 pp 197-199)) is richer than the PEM model of the mechanism of attention. The mechanism of attention (of all stripes) is indeed the variability of precision-weighting in every neuronal area and at every level of processing. But what determines how that variability is used is the task at hand, and information concerning internal and external context. Internal context will include interoceptive information concerning e.g. physiological indicators of hunger and thirst. But it will also include the effects of standing beliefs (here cast as predictions) operating at multiple time-scales. It is these multi time-scale beliefs that must now be doing much of the work (in concert with internal and external cues) in accounting for endogenous shifts of attention.

To bring this into focus, notice that these accounts make 'beliefs' (in the broadest sense) primary. This is already a common feature of many Bayesian accounts in which desired states of affairs are re-cast as beliefs that the desired state of affairs is observed, allowing inferential processes to uncover the ways to bring that state about (see e.g. Todorov (2009)). In the 'active inference' model, desires are simply beliefs/predictions that thus guide inference and action (see Friston et al (2011) p. 157). My desire to *drink a glass of water now* is cast as a prediction that *I am drinking a glass of water now* – a prediction that will yield streams of error signals that may be resolved by bringing the drinking about, thus making the world conform to my prediction. Desires are here re-cast as predictions apt to be made true by action.

Thus consider the prediction (based on some standing or newly emerging belief) that I will now experience, say, the hand-clapping film. This would enslave action, including the 'mental action' of altering the precision-weighting on hand-clappy stuff. In this way desires and motivations are revealed as beliefs that enslave action. The apparently non-indicative nature of a thought such as 'let's have a look at the hand-clap film' is now no barrier. For the real content of the thought, as far as the PEM mechanism is concerned, is indicative – it is something like 'I am looking at the hand-clap film now'. This informs our expectations of precision, creating the kind of partially self-fulfilling loop described earlier. High-level (initially false) predictions concerning which film is currently being viewed thus cause the precision re-allocations that enable that very film to switch into awareness.

5. The Bigger Picture

Ransom et al (p.110) do (very briefly) consider this kind of response. But they argue that such a move is inadequate, largely on the grounds of potential circularity. It is simply not clear, they suggest, how the move from desires to beliefs or policies can explain the occurrence or timing of voluntary shifts of attention. An 'action policy' that says 'attend to what you want', or 'attend to what you decide to attend to' would, they note, do the job - but only at the cost of evident circularity. What is still missing, they claim, is a satisfying PEM story about the *initiation* of voluntary action including voluntary shifts of attention.

This response fails to recognize the true scope of the formal demonstration that *any* set of behaviors prescribed by reward, cost, or utility functions can be prescribed by an apt set of systemic beliefs or priors (Friston (2011), Brown (1981), Friston et al (2012)). Given that result, it is clear enough that the behavioral profile associated with any set of desire or motivation-driven responses can be modeled as a set of responses driven by delicately interwoven webs of belief, driving predictions at many different time-scales.

To be sure, the question could still be pressed: just where do these self-predictions (that get to enslave action and mental action) come from? But this is clearly no worse than the parallel question: how do desires and motivations – such as a sudden desire to watch the other movie - arise and act in the brain? Whatever account Ransom et al favour of the latter can simply be incorporated into the PEM treatment, translating desire/utility talk into belief/prediction talk, as rehearsed above. There is no circularity here. Whatever set of personal and environmental circumstances might conspire to install or suddenly foreground a desire (for example, the desire to attend to a different movie channel), those same circumstances are now called upon to install or suddenly foreground the behaviorally-equivalent belief. That belief-complex includes bespoke precision expectations that then become self-fulfilling in the standard PEM manner, bringing about the alterations in precision weighting that control the flow of information in the brain, thus favouring one movie over the other.

To round this story off, notice that PEM already commands a potent mechanism capable of accounting for endogenously-driven change in systemic predictions. That mechanism is the joint (interoceptively *and* exteroceptively driven) minimization of prediction error (Pezzulo, Rigoli, and Friston (2015)). This ensures that complex organisms respond to an ever-changing internal milieu (potentially including some degree of random noise) in ways that dovetail with environmental opportunities. It is easy to see, for example, how an interoceptive prediction error signal created in response to a decrease in blood glucose might combine with perceptual information to entrain an opportunistic act of eating. Such an act, from the PEM perspective, is initiated not on the basis of a simple desire to eat, but on the basis of a standing belief that eating will occur when (for example) blood sugar levels are predicted to dip below a certain point. Once interoceptive information suggesting such a dip is received, there emerges a prediction (aimed at resolving that interoceptive prediction error) that *eating now occurs*.

Importantly, this shows that precision should not be thought of simply as encoding the reliability of a signal, but also its estimated *value (*Friston et al (2012)). The link is firm since high precision predictions will exert greater control over action, corresponding (informally) to greater motivational salience (see Friston (2009) p.300). Value (including the 'epistemic value' of exploratory actions) thus emerges as the complement of surprise, to be maximized by action (Pezzulo et al (2015) p. 33). Shifting precision assignments also result from perceived opportunities to achieve a goal, where that means: the opportunity to make some standing or newly minted prediction come true, at this very moment, by engaging in apt action. For an impressively comprehensive sketch of this PEM-based reconstruction of goal-directed behavior and motivational hierarchies, see Pezzulo et al (2015).

In the context of a deep generative model, motivational flux thus emerges naturally as changing environmental opportunities interact with changing internal states, both spawning and responding to self-predictions spanning multiple temporal scales. These labile self-predictions play the role of shifting patterns of desire, and give rise to the changing precision weightings that implement voluntary attention.

5. Conclusions: Broadening Precision

Ransom et al's careful critique helps reveal the pressing need to broaden our vision of precision estimation itself. For precision estimations must play multiple interlocking roles in a fully described PEM economy. Ransom et al focus mostly on one core role – that of tracking estimated reliability. They also allow that merely estimating a signal to have high-precision may sometimes create a positive feedback loop that helps bring that very state of affairs about (making a weak signal 'jump out' from the noise). But precision is adjusted (by multiple means) for every neuronal populations, and is thus entangled with many different kinds of information processing, including the ongoing interoceptively-informed estimation of both gross and epistemic value.

Within this evolving regime mental actions, such as voluntarily attending to one film rather than another, are indeed subtly special. Such mental actions are not the results of voluntarily altering our own precision estimations. Switches in what is attended cannot *result* from altering precision estimations because, just as Ransom et al argue, the switching *consists* in those very alterations. Instead, such mental actions are nothing other than the complex, allocations and re-allocations of precision themselves. But these re-allocations are not thus rendered mysterious or beyond the remit of the PEM-story. For they are themselves determined by our standing beliefs, evolving inner states, and the play of perceived external opportunities. The ghost of unanalyzed desire is thus banished from the predictive machine, replaced by intentions in the form of prior beliefs.

Acknowledgements

Thanks to the two anonymous referees, whose comments have improved the clarity and enhanced the scope of this response.

Funding: This paper was written thanks to support (in the form of sabbatical leave) from the School of Philosophy, Psychology, and Language Sciences at the University of Edinburgh, and ERC Advanced Grant XSPECT - DLV-692739.

References

Brown, L. (1981). A Complete Class Theorem for Statistical Problems with Finite Sample Spaces. *Ann. Stat*. 9, 1289–1300.

Brown H, Friston K, and Bestamnn S. (2011) Active inference, attention and motor preparation. *Frontiers in Psychology* 2: 218 (doi: 10.3389/fpsyg.2011.00218)

Brown, H., Adams, R. A., Parees, I., Edwards, M., and Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cogn. Process.*

Clark, A. (2013a). Whatever Next? Predictive brains, situated agents, and the future of cognitive science. *Beha vioral and Brain Sciences*, *36*, 181–204.

Clark, A. (2013b). The Many faces of Precision *Frontiers in Psychology*, *4*.

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind* (Oxford University Press, New York)

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–36.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7).

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–38

Friston, K. (2011). Perspective What Is Optimal about Motor Control ? *Neuron*, *72*(3), 488–498.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference, 137–160.

Friston, K., Samothrakis, S., & Montague, R. (2012). Active inference and agency: Optimal control without cost functions. *Biological Cybernetics*, 106, 523–541.

Hohwy, J. (2013). *The predictive mind.* Oxford University Press.

Kanai R, Komura Y, Shipp S, Friston K. (2015) Cerebral hierarchies: predictive processing, precision and the pulvinar. *Phil. Trans. R. Soc. B* 370: 20140169.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7(4), 480–494.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, 1–19.

Ransom, M., Fazelpour, S., & Mole, C. (2017). Attention in the predictive mind. *Consciousness and Cognition*, *47*, 99–112.

Todorov, E. (2009). Efficient computation of optimal actions. *PNAS*, 106(28).